University of Plymouth

PEARL **https://pearl.plymouth.ac.uk**

01 University of Plymouth Research Outputs University of Plymouth Research Outputs

2019-08-30

Learning from significant medical events: a systematic review

Price, Tristan

http://hdl.handle.net/10026.1/12423

10.5750/ejpch.v7i2.1637
European Journal for Person Centred Healthcare

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Learning from significant medical events: a systematic review

Abstract

Rationale, aims and objectives: Learning from significant medical events is a core component of quality and safety practice in healthcare worldwide. However, the evidence that analysis of, or reflection on, significant events has a positive impact on subsequent doctor performance is relatively sparse. This review aims to explore the impact of undertaking significant event analysis on medical performance.

Method: A systematic review using the following databases: PubMed, EMBASE, Medline, PsycINFO and the Cochrane Collaboration Library. Citation searches were carried out on included studies. Impact was defined according to a modified adaptation of the Kirkpatrick evaluation model. The selection and quality appraisal of studies was conducted by two reviewers, independently and blinded. Data were extracted from included studies related to: study type and location, population, methodology and intervention type.

Results: Six papers met the inclusion criteria for this review. Of these: one reported learners' reaction (Kirkpatrick 1); two reported modified attitudes (modified Kirkpatrick 2a); five reported the acquisition of knowledge (modified Kirkpatrick level 2b); and all six identified reported changes in behaviour (modified Kirkpatrick level 3a). Significant event analysis is reported to identify gaps in knowledge, improve teamwork and communication, and encourage reflection leading to improvements in practice. Time, resources and team dynamics were identified as factors that impacted on the effectiveness of significant event analysis. Significant event analysis may benefit from suspending existing hierarchies during the process itself, and external facilitation.

Conclusion: There is a lack of high quality evidence within the existing literature to ascertain the effectiveness of significant event analysis in the medical context. Existing studies are largely based on self-reported measures, which may reinforce the importance of the discursive process for practitioners. Future research could be directed at identifying the pedagogical processes that lead to changes in performance as a result of engaging in significant event analysis.

Introduction

At its core, person centred care is about providing healthcare services that take the needs and comfort of the patient as the primary concern and the focus for healthcare improvement. Ensuring patient safety, perhaps the most important aspect of person-centred care, is now a top priority for most modern healthcare systems,[1] but this has not always been the case despite increasing evidence of adverse events through the 1980s and 1990s.[2] However in 1999, the Institute of Medicine in the United States published what is widely regarded as a seminal report into healthcare failings; *To Err is Human* highlighted how preventable medical errors were a significant cause of death and injury, exceeding the numbers of deaths attributed to vehicle accidents, breast cancer and AIDS.[3] The report argued for nationwide systems for collating incidents and for clinicians to take a role in their development. The report created "an international sense of urgency to reduce harm in healthcare delivery",[2, p.1] and subsequently, incident reporting and analysis has been an increasingly prominent feature of quality and safety improvement systems, including throughout Europe.[4, 5]

Adapted from procedures that had been developed in the aviation industry,[6] the basic rationale behind significant event analysis (SEA) is that, through carefully examining key events – usually those that could have or did lead to significant harm – learning can take place at both an individual and an organisational level to improve quality and safety.[7] Pringle et al.'s 1995 paper provides a workable and useful definition of SEA for this review:

> '[SEA is where] individual cases in which there has been a significant occurrence…. are analysed in a systematic and detailed way to ascertain what can be learnt about the overall quality of care and to indicate changes that might lead to future improvements.'[8]

This inclusive but specific interpretation of what constitutes SEA allows us to capture the broad range of activities that happen in different healthcare sectors. Often in primary care the focus will be on the actions and responses of an individual doctor in regard to a consultation or series of consultations with a specific patient. In secondary care, a root cause analysis (RCA) may be undertaken with those involved in the event to ascertain the causes and the appropriate measures that need to be instigated. For the purposes of this review these processes are classified as SEA if they involve an individual doctor engaging with the process of reviewing a significant event for the purpose of identifying areas for improvement. This can be distinguished from audit whereby doctors compare more general practice (e.g. prescribing) against accepted standards within the field.

The existing literature has tended to focus on the impact of SEA on the function of organisations, using outcomes measures such as mortality rates. Stavropoulou et al., have provided a comprehensive overview of the literature. While the authors found some evidence of change to clinical settings or processes as a consequence of establishing incident reporting systems, there was little evidence for improved outcomes. [9]

However, regulatory developments, such as medical revalidation in the UK, are mandating that doctors provide evidence of engagement in SEA as part of their personal development, under the assumption that such engagement improves medical performance.[10, 11] Given such developments it is important to consider the current evidence base for the assumption that engaging in SEA improves doctor performance.

A further rationale for examining the evidence for the impact of SEA on doctor performance is that, despite being a well-established method of safety and quality improvement both in the UK and internationally [12, 13], SEA has also been subject to some recent criticism related to its role and function. Dodds and Kodate have argued that significant event analysis in the UK has undergone "institutional conversion", a process by which the original role of SEA as a formative procedure to engender learning has mutated to a summative function used to grade healthcare organisations.[14] Dodds and Kodate point to the way in which data collected on significant events and "never events" – a prescribed list of events that are deemed both serious and avoidable such that they should never happen in a healthcare setting – are being collated and processed by the same institutional bodies that undertake NHS commissioning. This institutional conversion has resulted in a situation in which significant event reporting becomes a performance indicator for healthcare providers, working against the educational ethos behind SEA in which transparent reporting of significant events is the first step in the process of reflective learning.

Amongst other criticisms, Macrae has argued that SEA requires infrastructure for routine and detailed investigation which, although well developed in the aviation industry, is not sufficiently developed in healthcare to deliver the same measures of success. As a result, in healthcare settings there is, according to Macrae, a tendency to prioritise quantity over quality in incident reporting. Like Dodds and Kodate, Macrae also points to the way in which the large volume of incident reporting is used a measure of safety rather than a process of learning. This creates a culture in which SEA becomes a "passive process", in which staff see incident reporting as merely a way of logging problems and waiting for fixes, removing responsibility for local improvement.[4] Others have pointed to the lack of evidence within the wider literature that SEA improves either practitioner performance or patient safety.[2, 15]

Given the ongoing debate on the efficacy of the SEA, and in response to broader regulatory changes within medicine that are increasingly seeking to formalise and quality assure systems for reporting and analysing significant events, this review aims to assess the impact of SEA by addressing the following research questions:

- RQ1: What impact does SEA have on doctor performance?
- RQ2: What factors influence the effectiveness of SEA in the medical context?

Methods

This review follows a systematic approach based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) [16] and the Centre for Reviews and Dissemination guidance.[17]

Search Strategy

The search strategy was developed using the PICO framework (see Table 1). This involved the development of a series of keywords and their synonyms related to the population, intervention and outcomes derived from the research questions. All search terms were reviewed by several members of the research team and assessed according to the Peer Review of Electronic Search Strategies (PRESS) guidance.[18]

Authors sought the advice of an information specialist to develop a search strategy and the following databases were searched: Cochrane, EMBASE, Medline, PubMed,

PsycInfo. Results were filtered for articles published after 2006 and in the English language.

Table 1: Systematic review search terms

Search strategy
Population: anesthesiologist OR anesthesiologists OR anaesthesiologist OR anaesthesiologists OR anaesthetist OR anaesthetists OR "general practitioner" OR "general practitioners" OR "family practitioner" OR "family practitioners" OR GP OR GPs OR obstetrician OR obstetricians OR gynaecologist OR gynaecologists OR gynecologist OR gynecologists OR paediatrician OR paediatricians OR pediatrician OR pediatricians OR ophthalmologist OR ophthalmologists OR pathologist OR pathologists OR cardiologist OR cardiologists OR dermatologist OR dermatologists OR gastro-enterologist OR gastro-enterologists OR immunologist OR immunologists OR oncologist OR oncologists OR rheumatologist OR rheumatologists OR psychiatrist OR psychiatrists OR radiologist OR radiologists OR urologist OR urologists OR surgeon OR surgeons OR doctor OR doctors OR physician OR physicians OR consultant OR consultants OR specialist OR specialists OR locum OR locums Or clinician OR clinicians OR clinical OR "general practice" OR "primary care" OR "secondary care" OR clinic OR surgery **AND** **Interest:** "serious untoward incidents" OR "significant event" OR "significant events" OR "critical incident" OR "critical incidents" OR "never events" OR "serious untoward incidents" OR "serious untoward events" **AND** **Outcome:** "professional development" OR change OR improve OR quality OR learning OR reflect OR reflection OR impact OR outcome OR safety OR competence OR effective OR performance

Study selection

In the first stage of screening two reviewers independently examined the titles and abstracts from all returned articles after duplicates had been removed. Abstract screening was conducted using *Rayyan*, [19] a web-based application specifically designed for collaborative systematic reviews. Using this software, both reviewers conducted the screening independently (blind) and then worked through any conflicts together. If agreement could not be reached from reading the title and abstract, then the study was included for full text screening.

During the second stage of screening the same two reviewers independently assessed full texts of the articles. Where agreement could not be reached, the article went to a third reviewer for input until a resolution could be achieved.

Inclusion criteria

In order to be included, studies had to meet the following criteria: published in the English language between the years 2006 and 2017; explore the impact of SEA on doctor performance; be of any study design except review articles (i.e. non-systematic reviews) or commentaries (Table 2).

Exclusion criteria

Studies that discussed the development of new tools or processes for analysis of significant events were only included if they revealed data on the impact of SEA more

generally. Given that the outcome measure for the review was medical performance, studies that focused on patient safety outcomes were only included if they linked directly to an assessment of the impact on medical performance; patient safety indictors were not taken as a proxy measure for medical performance as it was decided that any change in patient safety outcomes could just as likely be due to changes in protocols at the organisational level rather than specifically the performance of doctors. Finally, due to the setting of interest, studies in the context of undergraduate medical education were excluded (Table 2).

Table 2 Study inclusion criteria form

Inclusion criteria form
1. Is the study available in English? Yes (proceed) No (reject)
2. Is the study published between 2006 and 2017? Yes (proceed) No (reject)
3. Does the study report on the impact of SEA on medical performance? Yes (proceed) No (reject)
4. Is the study focused on SEA in the context of undergraduate medical education? Yes (reject) No(proceed)

Data extraction

Data extraction was undertaken by one reviewer using a piloted extraction form. This was checked by a second reviewer with any necessary amendments being made. Information extracted included: year published; study location, aim and design; study population and sample methodology; and intervention type.

Outcomes

The specific outcome measure for RQ1 of this review was medical performance. After discussion with the review team, Barr at al.'s adaption of Kirkpatrick's four-level evaluation model[20], which has been further modified by Overeem et al.[21], was used to evaluate all included study outcomes (Table 3). Decisions on which classification of Kirkpatrick's hierarchy of evaluation were taken by two reviewers working together. This framework allows the results to be understood in terms of the nature of evidence provided, allowing for a more coherent and precise synthesis of research findings. For RQ2, where studies had identified changes to doctor performance, we sought to identify any factors that facilitated or impeded this process.

Table 3: Kirkpatrick's four level evaluation model as modified by Barr et al. (2000)[22] and Overeem et al. (2010).[21]

Level	Description
Level 1: learners' reaction	Relates to participants' views of their learning experience programme
Level 2: learning outcomes	
2a: modification of attitudes/perception	Changes in reciprocal attitudes or perceptions between participant groups, towards patients/clients and their condition, circumstances, care and treatment.
2b: Acquisition of knowledge/skills	Acquisition of concepts, procedures and principles of inter-professional collaboration or the acquisition of thinking/problem-solving, psychomotor and social skills linked to collaboration
Level 3: Change in behaviour	Behavioural change transferred from the learning environment to the workplace prompted by modifications in attitudes or perceptions, or the application of newly acquired knowledge/skills in practice.
3a: Self-reported change in behaviour	
3b: Measured change in performance	
Level 4: Patient/Organisational outcomes	
4a: Change in organisational practice	This relates to wider changes in the organisation/delivery of care, attributable to an education programme.
4b: Benefits to patients/clients	Covers any improvements in the health and well-being of patients/clients as a direct result of an education programme.

Quality assessment

Quality assessment was undertaken independently by two reviewers. The Critical Skills Appraisal Programme (CASP) checklists for qualitative studies[23] and systematic reviews[24] were used for quality assessment. Quality assessment was important to assess the overall strength of findings for the review, but articles were not excluded on the grounds of their quality as conceptual relevance was deemed more important, especially given the limited literature available on this topic.

Data analysis and synthesis

We followed Popay et al.'s modified narrative synthesis technique. This process of narrative synthesis (which can be distinguished from more general and non-systematic "narrative reviews"), is based on the premise that systematic synthesis of qualitative evidence can be presented in such a way that it tells a "trustworthy story" about the nature of evidence within the field of study. To achieve this, researchers develop a framework for analysis, undertake a preliminary synthesis, explore relationships between the data and assess – through quality appraisal – the robustness of the findings. Importantly, this process is not always sequential and researchers may, through an iterative process of collaboration, modify and adjust the parameters of their research, and even the research questions, as they learn more about the scope and nature of the available evidence.[25]

Following these broad guidelines, we used accepted methods of thematic synthesis to analyse the data from the included studies,[26, 27] using a combined (inductive and deductive) approach.[18] An overarching coding framework based on the two aspects of the research questions (the impact of SEA and factors influencing its impact) provided an initial scaffold of parent codes. Subsequently, an inductive approach was used to further populate the framework with more nuanced categories identified from the data. Through an iterative and collaborative process between two reviewers, the coding framework was modified as data was extracted and themes emerged. For RQ1, once the data was coded it was then mapped onto the adapted Kirkpatrick framework to determine the level at which change was taking place.

Results

Six studies met the inclusion criteria. The full process of study selection is provided in Figure 1.

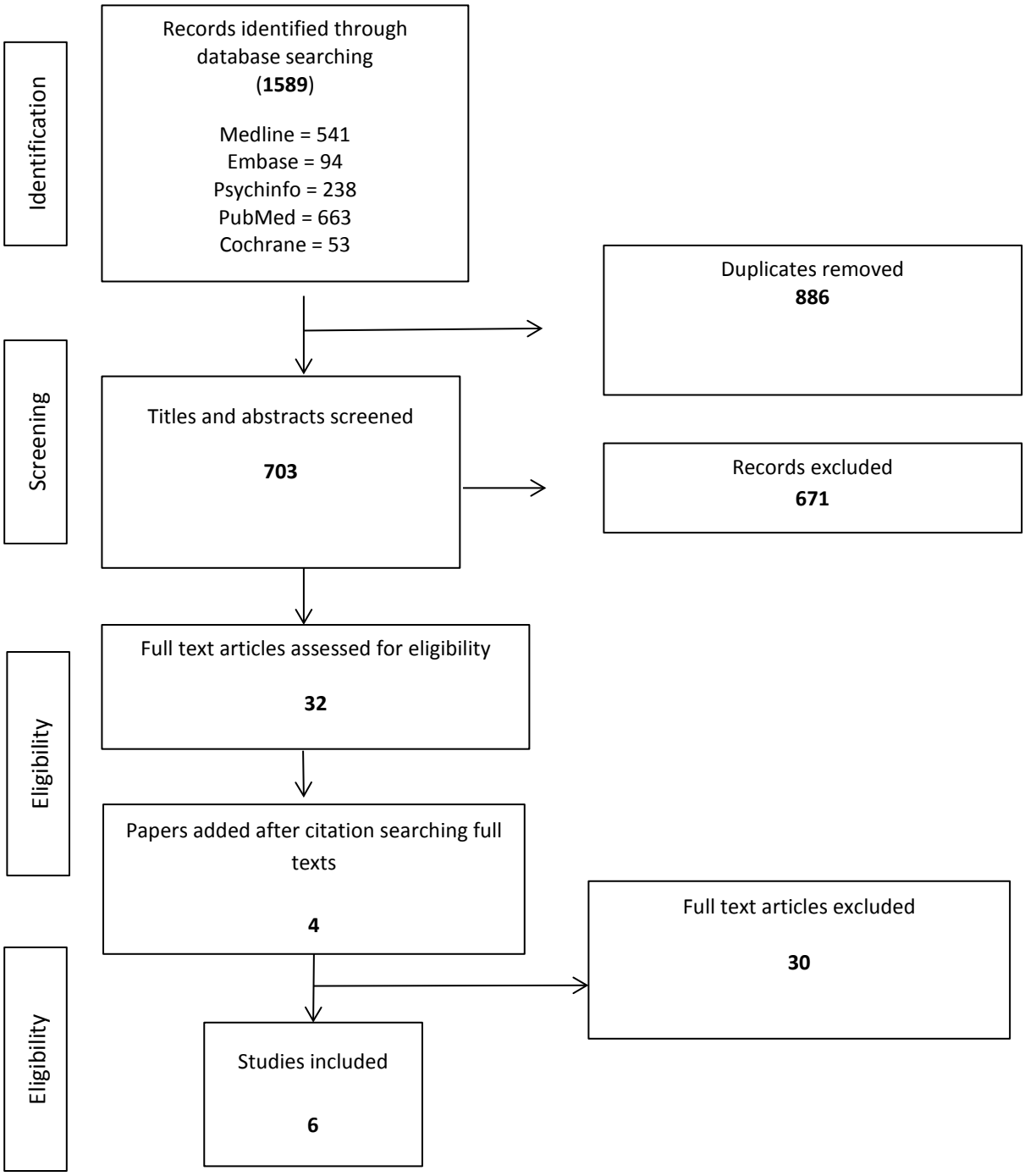
Study Characteristics

Of the six studies that met the inclusion criteria, three related to primary care,[28-30] two to secondary care[6, 31] and one which covered both sectors.[7] The included studies in secondary care both focused on root cause analysis (RCA). One study was a systematic review of the published literature up until 2006. This review reported on studies in both primary and secondary care, although the majority of studies reported were from primary care.[7] Of the five original studies included: one was a review of significant event reports,[29] two used questionnaire surveys of, or including, practising doctors,[28] and two were based on semi-structured interviews with, or including, practising doctors.[30, 31]

Given that SEA is a well-embedded system to assure the quality and safety of healthcare organisations,[4, 5] especially in secondary care, the small number of included studies is in itself worthy of comment. A high number of studies in both primary care and secondary care were excluded because they focused on the organisational level, usually using patient outcome data to evaluate the impact of incident reporting systems. This literature has been comprehensively reviewed elsewhere.[9] Other excluded studies focused on the development of new procedures for incident reporting[13] or sought to assess the impact of a training programme to improve incident reporting processes,[32-34] but without sufficiently focusing on medical performance outcomes to meet our inclusion criteria. Some other studies, in both primary and secondary care, were expressly focused on barriers to effective incident

reporting processes, but without examining any change in medical performance resulting from engagement with those processes.[35, 36]

Figure 1: PRISMA flow diagram



In summary, only a small number of articles met our inclusion criteria because the vast majority of recent studies on SEA do not focus on the impact of participation in SEA on doctors' performance.

RQ1: Impact of Significant Event Analysis on Doctor Performance

Out of the six included studies: one reported learners' reactions[7] (Kirkpatrick 1); two reported modified attitudes[7, 31] (modified Kirkpatrick 2a); five reported the acquisition

of knowledge[6, 7, 29-31] (modified Kirkpatrick level 2b) and all six identified reported changes in behaviour [6, 7, 28-31] (modified Kirkpatrick level 3a). No studies reported measured changes in behaviour (modified Kirkpatrick level 3b), changes to organisational practice (modified Kirkpatrick level 4a), or outcomes related to patient safety (modified Kirkpatrick level 4b). It is unsurprising that no studies reported changes to organisational practice (level 4), given that we excluded studies that used patient outcome metrics to measure changes to organisational performance.

Applying the modified Kirkpatrick evaluation framework was not always straightforward. One of the issues that emerged was that some studies reported organisational changes (potentially modified Kirkpatrick level 4a), but this change was self-reported as opposed to measured, i.e. doctors involved in the study reported their belief that organisational change had occurred. After a closer reading of the modified Kirkpatrick framework and subsequent discussion between reviewers (TP and RB), it was agreed that such reporting constituted a level 3a and that outcomes would have to be measured to qualify for a level 4.

In light of this, we classified the systematic review as 1, 2a, 2b and 3a, and not level 4, as the evidence reviewed, in the main, came from studies based on the identification of learning needs and self-reported change. Where the review discussed changes to patient care and service delivery (level 4), it suggested that the studies from which this was drawn were based on self-reported measures, e.g. “SEA is *reported* to have facilitated multiple changes in practice and important improvements in service quality”.[7, p.523]

Learners’ reaction (Kirkpatrick Level 1)

This level of Kirkpatrick framework is for those findings that relate to personal reactions to engaging in SEA but in situations whereby those reactions are not directly linked to improved learning or behaviour. Only the systematic review by Bowie et al. reported findings on this level. On the one hand, the review found that engaging with SEA, provided “an element of personal catharsis…for some, especially in writing up the report”. On the other hand, they also found that it could engender reactions related to, “a fear of litigation, exposure, reprisal or humiliation”. The “critical process” of SEA was reported as being sometimes “disconcerting”, which could “release un-containable pressures”.[7, p.523]

Modification of attitudes / perception (modified Kirkpatrick Level 2a)

Studies reporting on changing attitudes as a result of SEA described such changes in quite general terms. One study into the use of RCA in Iranian hospitals found that RCA could lead to an “attitude change among healthcare workers towards safer improvement strategies”.[31, p.157] Similarly, the systematic review into SEA found that it could lead to “a greater commitment to change” amongst and could “generate greater confidence and higher levels of personal trust” amongst staff.[7, p.523]

Acquisition of skills / knowledge (modified Kirkpatrick Level 2b)

Engaging in SEA was reported as identifying gaps in knowledge and learning needs of both individual practitioners[7, 29, 30] and teams.[31] 95.3% (n=182) of the significant event reports analysed by McKay et al. found that learning needs were identified in the reports themselves. These related to: general awareness of personal issues, responsibilities and change (51.3%, n=98), clinical knowledge (15.7%, n=30), the use of

equipment and workspace (13.6%, n=26), and medication/prescribing (4.7%, n=9).[29, p.67]

General Practitioner participants in the study by Rea and Griffiths had commented on the "emotive value" of SEA that "enabled them to personally reflect on their mistakes" and that "learning from deviations in care" was the "over-riding purpose for reporting and analysing incidents".[30, p.5] Similar findings were reported in secondary care with 81% (n=36) of respondents in Abdi and Ravaghi's study into RCA in Iranian hospitals "believed that the process of going to the root of a problem", enabled doctors to "develop a new understanding" and to "challenge taken-for-granted assumptions".[31, p.156]

Self-reported changes in behaviour (modified Kirkpatrick level 3a)

Studies gave mixed results on self-reported changes to behaviour as a result of engaging in SEA, although all recognised some positive change.

Three studies reported changes to clinical practice.[7, 28, 29] SEA was reported to have "facilitated multiple changes to practice and important improvements in service quality".[7, p.523] 91% (n=29) of the participants in the study by Abdi and Ravaghi described RCA as a "valuable tool for improving work practices".[31, p.156] The study by McKay et al., which analysed significant event reports in primary care delivered similar findings, with 81% of those reports (n=154) demonstrating that "change(s) had been agreed and implemented by at least one member of the primary care team as a result of SEA". These changes related to disease diagnosis and management, (16.6, n=32), doctors' personal skills/behaviour/knowledge application (14.6%, n=28), and communication (13.6%, n=26).[29, p.65] This is supported by a similar study by de Wet et al., in which 85% of respondents (n=93) reported that engaging in SEA had resulted in a change in practice.[28, p.1210]

In addition to the McKay et al. study, three of the other included studies specifically reported that SEA had improved teamwork and communication.[6, 7, 31] Participants in the study by Abdi and Ravaghi expressed a "common belief" that root cause analysis "provides an opportunity to improve teamwork and communication by bringing staff from different disciplines together to discuss system vulnerabilities".[31, p.156] This finding was supported in the study by Braithwaite, who found that 80.3% (n=200) of respondents reported that RCA improved teamwork, and 79.8% (n=201) that it improved communication about patient care.[6, p.395]

Similarly, Bowie et al. describe the role of SEA in bringing together the "the wider team to contribute to decision making when resolving problems and conflicts, which may lead to staff feeling valued and successful change being sustained." In particular, the authors argue that a "focus on team dynamics may encourage a shift from an individualistic view of practice to a broader teams-based perspective".[7, p.522]

Other findings related to the role of SEA in providing an opportunity for both individual [29, 30] and team reflection.[7] Similarly, the study by McKay et al. found "further evidence" that SEA "provided an opportunity for reflection", but warned that, "much of the learning would appear to be personal to the SEA report author" rather than the wider clinical team.[29, p.70] In contrast, the systematic review by Bowie et al. reported evidence that SEA provided, "an opportunity for the team to reflect that may generate greater confidence and high levels of personal trust".[7, p.522] In considering the difference in findings between these two studies, it is worth noting that the methodology

for the McKay et al. study was a content analysis of SEA reports, so it is difficult to see how this methodology could have delivered any robust findings relating to the reflection of anyone other than the report author.

Although, as noted above, some studies had reported negative emotional reactions to engaging with SEA, none of the included studies reported negative changes to behaviour e.g. defensive clinical practice.

RQ2: Factors that influence the impact of significant event analysis on medical performance: Team dynamics, leadership and resources

Four of the seven included studies reported on the factors that influenced the impact of SEA on doctor performance.[6, 7, 30, 31] All four of these reported team dynamics and leadership as a prominent factor. This was the case in both primary care[7, 30] and secondary care (RCA).[6, 31]

Established and cohesive teams can aid in the process of SEA if there is a "high degree of trust and openness" and a "supportive and open working environment" to facilitate open discussion. [7, p.524] However, SEA may not always be conducted with the whole team, thus hindering team reflection.[30] Poor team dynamics may also "militate against the critical appraisal of the care delivered".[7, p.523] In particular, poor team dynamics can be attributed to existing hierarchies whereby more junior team members may be reluctant to challenge the interpretation of events proffered by their senior colleagues.[7, 31] This in turn may lead to a situation of "group think", whereby a team consensus is arrived at, not because it is the most valid interpretation of events, but "because of the number or seniority of those that expressed it".[31, p.154] As a result, the process of SEA should be one in which hierarchy is suspended,[7] should include both clinical and non-clinical staff,[31] and could benefit from regular meetings and external facilitation, provided any such facilitator is appropriately trained.[7]

An equally prominent factor in the included studies was that of time and resources. In the case of RCA, 78% (n=25) of participants in the study by Abdi and Ravaghi expressed that gathering sufficient information was the most challenging aspect of conducting RCA.[31, p.153] This finding was supported by Braithwaite et al., who found that 75% (n=189) of respondents stated that a lack of sufficient time allocated to RCA hindered its effectiveness and "always or sometimes occurred."[6, p.395] This is often coupled with a more general lack of resources that hinders not only the investigation into significant events,[6] but also the implementation of resulting strategies.[31]

This was also the case in primary care, where the "busyness of general practice" and the priority accorded clinical work sometimes means that insufficient time is available for SEA.[30, p.5] This in turn can lead to a lack of continuity because meetings to discuss significant events do not always have the relevant staff members in attendance.[7]

Other factors impacting on the effectiveness of SEA include a fear of blame or litigation, which may prevent doctors from reporting events in the first place,[30] and the need for effective training in both identifying and reporting significant events.[7]

Six out of nine GPs interviewed for the study by Rea and Griffiths stated that the requirement to reflect on significant events within their appraisal motivated doctors to produce reports and that the appraisal process itself encouraged a culture of openness around significant event reporting.[30]

Quality appraisal

The studies were of generally medium quality, with the systematic review being low quality due to a lack of rigour in reporting the search protocols. Moreover, even where one study was of a higher quality, the outcome measure of doctor performance was not its main focus. Accordingly, the conclusions are drawn from a relatively weak evidence base.

Summary of findings and Discussion

- *RQ1: What impact does SEA have on doctor performance?*

In 2008 Bowie et al. described a "faith-based assumption" that "engagement in SEA has positive benefits". SEA was popular, not because it is demonstrably efficacious in terms of improving doctor performance, but because it "is educationally attractive with high face validity".[7, pp.524-525] The results of this review suggest that this conclusion may still be valid. The low number of studies that met the inclusion criteria suggest that research into SEA has not generally been focused on its impact on doctor performance. The lack of high quality studies adds further weight to the argument that it is an under-researched area in the evaluation of clinical practice.

In categorising evidence according to a modified Kirkpatrick framework, we can also see that the changes in medical performance as a result of engaging in SEA are self-reported rather than measured, indicating that there is also scope for developing research methodologies that would seek to use some kind of performance indicators to objectively measure the impact of SEA. However, it should also be recognised that this would not be a simple endeavour. SEA does not lend itself to the development of the kind of proxy measures of performance that can, and have been used effectively to measure the effects of other interventions. In the case of audit and feedback for example, prescribing levels or guideline adherence can be used to ascertain changes to professional practice.[37] Whereas in the case of SEA, there are no obvious measures of performance, or proxy measures, that could easily distinguish between organisational and individual level outcomes.

There is also a lack of evidence in the literature relating to *how* SEA works. The *assumption* that SEA improves performance, which is common in the wider literature, is related to the idea that it improves reflection[8, 29] and that its benefits are, for example, "probably linked to adult learning theory."[38, p.735] Yet this review shows that there is a lack of theory-driven research into how SEA actually works to produce its effects, and what these effects are in terms of clearly measured outcomes of medical performance.

However, our findings also do not provide any evidence to refute the hypothesis that SEA improves doctor performance. The fact that the included studies highlighted self-reported changes to behaviour support the idea that those working in the field of medicine value the process of reflecting on incidents that occur in the workplace, which is supported more generally in the literature.[39, 40] Doctors also perceive SEA to be useful in highlighting deficiencies in skills, knowledge and improving teamwork and interpersonal skills. This evidence should not be discounted, but at the same time we should recognise that there is scope for more robust research into how SEA works and how it affects performance. Given the difficulties, described above, in determining clear and appropriate outcomes measures for SEA, future research may derive its robustness from qualitative methodologies and through linking research findings to theories of learning and pedagogy.

- *RQ2: What factors influence the effectiveness of SEA in the medical context?*

Where SEA is shown to be valued by those that engage in it, it is most effective when the process is accorded sufficient time and resources, and when there are good team dynamics and effective leadership. However, it is not clear from these studies what kind of resources are most critical, or exactly what effective leadership looks like. Moreover, if SEA becomes a more resource intensive process, it runs the risk of deterring practitioners from engaging with it. More research into best practice in SEA would therefore be useful.

The identification of team dynamics and leadership as facilitating SEA also poses a practical challenge in understanding how SEA works. Those teams with better dynamics and effective leadership will likely have less significant events in the first place and, correspondingly, poor team dynamics and leadership might be identified, through an effective SEA process, as being a cause of the events in question.[13]

A further challenge to understanding SEA is that the term itself may mean different things in different sectors or organisations. In the UK, SEA in primary care (family medicine) is more likely to be conducted by an individual practitioner and may include not only adverse incidents in which patient safety is compromised, but also positive outcomes. Whereas in secondary care, SEA is more likely to be conducted in teams, often as a RCA, and is focused specifically on adverse events that did or could have compromised patient safety. It may be that one model does not fit different sectors, but greater terminological clarity in terms of the process undertaken would aid in the evaluation of effectiveness.

SEA has also been identified as being dependent on effective processes of incident investigation that require a suspension of hierarchy in which participants have an equal input, regardless of their position within the organisation, and an atmosphere in which participants do not fear that they will attract blame when highlighting their own mistakes. This is particularly relevant in the current policy context in the UK, with the suggestion that a culture of blame and fear is pervasive in NHS institutions.[41]

This also speaks to some of the concerns that have been raised in relation to significant incident reporting in the UK whereby reporting systems have been accused of being subverted from their educational purpose in order to provide performance metrics for healthcare organisations.[4, 14] Our review suggests that, if SEA improves performance, it does so because, as Leistikow and colleagues put it, the “journey, not the arrival, matters”.[12, p.1] In particular, while individual reflection within performance appraisal may provide an incentive to engage in SEA, the evidence suggests it is the discursive process, within and between members of a team, which is most valued by participants and perceived to change behaviour. This supports the argument posited by Vincent, that linking SEA to root cause analysis is unhelpful because the latter is a “gross oversimplification” of the process. There is rarely, argues Vincent, a single cause of an event, so if SEA is to be effective it should be a “window on the system” not a retrospective search for root causes.[42]

This argument resonates with some of the wider literature on organisational development which warns against processes that focus solely on diagnostics at the expense of dialogue.[43] In terms of policy implications, these findings would therefore support efforts to develop more systems-focused SEA processes such as the Enhanced

Significant Event Analysis programmes that have been trialled by NHS Education for Scotland.[44]

However, if the aim of SEA is to create an open conversation about incidents that occur in healthcare organisations, then there is a difficult balance to be struck: measuring the performance of healthcare organisations in terms of incidents that occur may be important in highlighting institutional failings, but may in turn work against the development of an open culture in which practitioners candidly report and reflect on mistakes. This dilemma is important and somewhat paradoxical when viewed from a person-centred care perspective. As a means of ensuring patient safety, it may seem reasonable to make a judgement about the overall effectiveness of a healthcare provider by examining data on the significant events that occur within that organisation. However, this may undermine the capacity of SEA to induce improved performance by individual doctors by dis-incentivising full and candid engagement with the process, with potentially negative consequences for patients.

Strengths and limitations

This review benefits from a rigorous systematic methodology. The modified Kirkpatrick framework has enabled us to delineate the nature of the evidence provided in the literature. The main limitation of this review is the lack of high quality studies and the absence of clear and objective outcome measures to ascertain impact on medical performance. However, this is mitigated somewhat by the quality appraisal and the use of the Kirkpatrick framework as we have been able to evaluate the strengths of the claims we are making. Limiting the search to English language literature, while a practical necessity, may mean that we have missed some relevant studies.

Conclusion

There is a lack of high quality evidence within the existing literature to ascertain the effectiveness of SEA to improve doctor performance. The existing evidence is largely based on self-reported measures, suggesting that it is a process valued by those that engage in it, but that its efficacy is based on assumptions rather than evidence. Benefits are reported as relating to improved teamwork and interpersonal skills and the provision of opportunities for reflection. SEA is perceived to be most effective when there are good team dynamics and when it is accorded sufficient time and resources. Effective leadership is deemed important but hierarchy is unhelpful in the process of investigating the causes of significant events. Future research could be directed at developing a more comprehensive and theoretically informed understanding of the pedagogical processes that lead to changes in performance as a result of engaging in SEA.

Acknowledgements: We would like to thank all those who contributed to the wider UMbRELLA research collaboration from which this paper arises.

References

1. McKay J, Bowie P, Murray L, Lough M. Levels of agreement on the grading, analysis and reporting of significant events by general practitioners: a cross-sectional study. Quality and Safety in Health Care. 2008;17(5):339-45.

2. Mitchell I, Schuster A, Smith K, Pronovost P, Wu A. Patient safety reporting: a qualitative study of thoughts and perceptions of experts 15 years after 'To Err is Human'. BMJ Qual Saf. 2015:bmjqs-2015-004405.

3. Kohn L, Corrigan J, Donaldson M. To err is human: building a safer healthcare system. United States; 1999.

4. Macrae C. The problem with incident reporting. BMJ Qual Saf. 2015:bmjqs-2015-004732.

5. Doupi P. National reporting systems for patient safety incidents: a review of the situation in Europe. Raportti/Terveyden ja hyvinvoinnin laitos (THL)= Report/National Institute for Health and Welfare: 13/2009. 2009.

6. Braithwaite J, Westbrook MT, Mallock NA, Travaglia JF, Iedema RA. Experiences of health professionals who conducted root cause analyses after undergoing a safety improvement programme. Quality and Safety in Health Care. 2006;15(6):393-9.

7. Bowie P, Pope L, Lough M. A review of the current evidence base for significant event analysis. Journal of evaluation in clinical practice. 2008;14(4):520-36.

8. Pringle M, Bradley C, Carmichael C, Wallis H, Moore A. Significant event auditing. A study of the feasibility and potential of case-based auditing in primary medical care. Occasional paper (Royal College of General Practitioners). 1995(70):i.

9. Stavropoulou C, Doherty C, Tosey P. How Effective Are Incident‐Reporting Systems for Improving Patient Safety? A Systematic Literature Review. The Milbank Quarterly. 2015;93(4):826-66.

10. Archer J, de Bere SR. The United Kingdom's experience with and future plans for revalidation. Journal of Continuing Education in the Health Professions. 2013;33:S48-53.

11. Archer J, Pitt R, Nunn S, Regan de Bere S. The evidence and options for revalidation in the Australian context. Medical Board of Australia 2015.

12. Leistikow I, Mulder S, Vesseur J, Robben P. Learning from incidents in healthcare: the journey, not the arrival, matters. BMJ Qual Saf. 2016:bmjqs-2015-004853.

13. Bowie P, McNaughton E, Bruce D, Holly D, Forrest E, Macleod M, et al. Enhancing the Effectiveness of Significant Event Analysis: Exploring Personal Impact and Applying Systems Thinking in Primary Care. The Journal of Continuing Education in the Health Professions. 2016;36(3):195-205.

14. Dodds A, Kodate N. Understanding institutional conversion: the case of the National Reporting and Learning System. Journal of Public Policy. 2012;32(2):117-39.

15. Mahajan R. Critical incident reporting and learning. British journal of anaesthesia. 2010;105(1):69-75.

16. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine. 2009;151(4):264-9.

17. Khan KS, Ter Riet G, Glanville J, Sowden AJ, Kleijnen J. Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews: NHS Centre for Reviews and Dissemination; 2001.

18. Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. Journal of clinical epidemiology. 2009;62(9):944-52.

19. Mourad Ouzzani, Hossam Hammady, Zbys Fedorowicz, Elmagarmid A. Rayyan — a web and mobile app for systematic reviews 2016 [Available from: https://rayyan.qcri.org/.

20. Barr H, Freeth D, Hammick M, Koppel I, S R. Evaluations of Interprofessional Education: A United Kingdom Review of Health and Social Care. London: CAPE; 2000.

21. Overeem K, Lombarts M, Arah OA, Klazinga NS, Grol RP, Wollersheim HC. Three methods of multi-source feedback compared: a plea for narrative comments and coworkers' perspectives. Medical teacher. 2010;32(2):141-7.

22. Barr H, Freeth D, Hammick M, Koppel I, Reeves S. Evaluations of interprofessional education. London: United Kingdom Review of Health and Social Care. 2000.

23. Critical Appraisal Skills Programme CASP (Qualitative Research) Checklist. Online 2017.

24. Critical Appraisal Skills Programme CASP (Systematic Review) Checklist. Online 2017.

25. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme Version. 2006;1:b92.

26. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. International Journal of Qualitative Methods. 2006;5(1):80-92.

27. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology. 2008;8(1):1-10.

28. de Wet C, Bradley N, Bowie P. Significant event analysis: a comparative study of knowledge, process and attitudes in primary care. Journal Of Evaluation In Clinical Practice. 2011;17(6):1207-15.

29. McKay J, Bradley N, Lough M, Bowie P. A review of significant events analysed in general practice: implications for the quality and safety of patient care. BMC family practice. 2009;10(1):61.

30. Rea D, Griffiths S. Patient safety in primary care: incident reporting and significant event reviews in British general practice. Health & social care in the community. 2016;24(4):411-9.

31. Abdi Z, Ravaghi H. Implementing root cause analysis in Iranian hospitals: challenges and benefits. The International journal of health planning and management. 2017;32(2):147-62.

32. Kantelhardt P, Müller M, Giese A, Rohde V, Kantelhardt S. Implementation of a critical incident reporting system in a neurosurgical department. Central European Neurosurgery. 2011;72(01):15-21.

33. Wallace LM, Boxall M, Spurgeon P, Barwell F. Organizational interventions to promote risk management in primary care: the experience in Warwickshire, England. Health services management research. 2007;20(2):84-93.

34. Bowie P, Skinner J, de Wet C. Training health care professionals in root cause analysis: a cross-sectional study of post-training experiences, benefits and attitudes. BMC health services research. 2013;13(1):50.

35. Nicolini D, Waring1 J, Mengis2 J. The challenges of undertaking root cause analysis in health care: a qualitative study. Journal of health services research & policy. 2011;16(1_suppl):34-41.

36. Elder NC, Graham D, Brandt E, Hickner J. Barriers and motivators for making error reports from family medicine offices: a report from the American Academy of Family Physicians National Research Network (AAFP NRN). The Journal of the American Board of Family Medicine. 2007;20(2):115-23.

37. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard‐Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. The Cochrane Library. 2012.

38. Cox SJ, Holden JD. A retrospective review of significant events reported in one district in 2004–2005. Br J Gen Pract. 2007;57(542):732-6.

39. Bruster BG, Peterson BR. Using critical incidents in teaching to promote reflective practice. Reflective Practice. 2013;14(2):170-82.

40. Bolsin S, Colson M, Patrick A, Creati B, Bent P. Critical incident reporting and learning. British journal of anaesthesia. 2010;105(5):698-.

41. Adrian R. The UK general election: a manifesto for health. The Lancet. 2015;385(March).

42. Vincent CA. Analysis of clinical incidents: a window on the system not a search for root causes. Quality and Safety in Health Care. 2004;13(4):242.

43. Bushe GR, Marshak RJ. Revisioning Organization Development: Diagnostic and Dialogic Premises and Patterns of Practice. The Journal of Applied Behavioral Science. 2009;45(3):348-68.

44. McNab D, McKay J, Bowie P. Evaluation of a Primary Care Safety Improvement Intervention using Enhanced Significant Event Analysis in a Regional Scottish Health Board. Quality in Primary Care. 2017;25(3):148-56.